

SECTION I: ASSESSMENT DEVELOPMENT

CHAPTER 2—OVERVIEW OF TEST DESIGN

LEARNING RESULTS

MEA questions are directly linked to the **content standards** and **performance indicators** described in Maine's *Learning Results*. The content standards are the basis for the reporting categories developed for each subject area; the performance indicators are used to help guide the development of test questions. No other content or process is subject to statewide assessment. An item may address part of, all of, or several of the performance indicators.

ITEM TYPES

Maine's educators and students were familiar with most of the question types that were used in the new assessment program; although one new type—the extended-response question—was used as well. The types of questions used and the functions of each are described below.

Multiple-choice questions were used, in part, to provide breadth of coverage of a subject area. Because they require no more than a minute for most students to answer, these questions make efficient use of limited testing time and allow coverage of a wide range of knowledge and skills.

Short-answer questions were used to assess students' skills and their abilities to work with brief, well-structured problems that had one or a very limited number of solutions (e.g., mathematical computations).

Short-answer questions require approximately two to five minutes for most students to answer. The advantage of this type of question is that it requires students to demonstrate knowledge and skills by generating, rather than merely selecting, an answer.

Constructed-response questions are the same as the “open-response questions” that have been used in past years of the MEA. These questions typically require students to use higher-order thinking skills—evaluation, analysis, and summarization, and so on—in constructing a satisfactory response. Constructed-response questions should take most students approximately five to ten minutes to complete. It should be noted that the use of previously released MEA questions to prepare students to answer this kind of question was appropriate and encouraged.

Extended-response questions were a type of question that had not been used previously in the MEA. These questions assess students’ ability to analyze and solve challenging problems based on real-world, age-appropriate situations that call for multiple approaches and may have more than one solution. An ability to communicate and justify a solution through the use of writing, tables, charts, and/or graphic displays contributes to a student’s success in many of the extended-response questions. This type of question requires approximately ten to twenty minutes for most students to complete.

COMMON-MATRIX DESIGN

In 1998–99, the MEA measured what students know and are able to do by using a greater variety of question types than it had used in the past. The tests continued to be structured using both *common* and *matrix-sampled* questions. Common questions are those taken by all students at a given grade level; in addition, a larger pool of matrix-sampled questions is divided among the multiple forms of the test at each grade level. (There were 16 forms of the test in 1998–1999.) Each student took only one form of the test and so answers a fraction of the matrix-sampled questions in the entire pool. This design, which has been used throughout the MEA’s history, provides reliable and valid results at the student level. It also provides for a greater breadth of coverage of a subject area for school results while minimizing testing time through the use of matrix-sampled questions.

In addition, each form of the test included a small number of questions that were being pre-tested, or field tested, using a statewide sampling of students. The pre-test questions were included to gather reliable statistics on question performance for possible future use in the MEA; students were not able to distinguish the pre-test questions from others in the test booklet.

There was a decision made by the Maine Department of Education to only report out common scores in the results. For the ease of understanding of results—in particular if student results based on common and matrix sampled, students could score higher than another in raw score, but lower in scaled score. Giving common only, this type of reversal was avoided.

TEST SESSION TIMES

The MEA tests were given at two different times during the school year: **English language arts, part 1** (writing) was administered to all grades in early November; tests in **all other subject areas** were administered to all grades during a four-week period extending from mid-March through early April. Schools were able to schedule testing sessions at any time during the first three weeks of this period, provided they followed the sequence in the scheduling guidelines detailed in test administration manuals. The fourth week was reserved for make-up testing of students who were absent from initial test sessions.

The tests given in the spring were divided into two units:

Unit 1—English language arts, part 2 (reading), science, and visual and performing arts;

Unit 2—mathematics, social studies, and health.

The test was divided into two units for greater ease of test administration and to minimize the impact of frustration a student might experience if the entire test booklet were one long document.

The timing and scheduling guidelines for MEA tests were based on estimates of the time it would take an average student to respond to each type of question that makes up the test:

- multiple-choice questions – 1 minute per question;
- short-answer questions – 2 minutes per question;
- constructed-response questions – 10 minutes per question;
- extended-response questions – 20 minutes per question (grades 8 and 11).

For the English language arts reading test, the scheduling guidelines included an estimate of 10 minutes to read each passage used in the assessment.

While the guidelines for scheduling are based on the assumption that most students will complete the test within the time estimated, each test session was scheduled so that additional time was provided for students who needed it. One-third additional time was allocated for each session (i.e., 60-minute sessions were scheduled with an additional 20 minutes; 45-minute sessions with an additional 15 minutes; and 30-minute sessions with an additional 10 minutes).

If additional classroom space was not available for students who required additional time to complete the tests, schools were allowed to consider using another space, such as the guidance office, for this purpose. If additional areas were not available, it was recommended that each classroom being used for test administration be scheduled for the maximum amount of time. Detailed instructions on test administration and scheduling were provided in the Coordinator's and Administrator's Manuals.